

Análise de medidas repetidas na avaliação de clones de café 'Conilon'

Paulo Roberto Cecon⁽¹⁾, Fabyano Fonseca e Silva⁽¹⁾, Adésio Ferreira⁽²⁾, Romário Gava Ferrão⁽³⁾, Antônio Policarpo Souza Carneiro⁽¹⁾, Edenio Detmann⁽¹⁾, Priscila Neves Faria⁽¹⁾ e Telma Suely da Silva Moraes⁽¹⁾

⁽¹⁾Universidade Federal de Viçosa (UFV), Centro de Ciências Exatas e Tecnológicas, Departamento de Informática, Avenida P.H. Rolfs, s/nº, Campus Universitário CEP 36570-000 Viçosa, MG. E-mail: cecon@dpi.ufv.br, fabyano@dpi.ufv.br, policarpo@dpi.ufv.br, detmann@ufv.br, prineves@vicosa.ufv.br, tel_morais@vicosa.ufv.br ⁽²⁾UFV, Campus Rio Paranaíba, BR 354, Km 310, CEP 38810-000 Rio Paranaíba, MG. E-mail: adesio@vicosa.ufv.br ⁽³⁾Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural, Rua Afonso Sarlo, nº 160, Bento Ferreira, CEP 29052-010 Vitória, ES. E-mail: romario@incaper.es.gov.br

Resumo – O objetivo deste trabalho foi analisar a produção de 50 clones de café 'Conilon' (*Coffea canephora*), por meio da análise de medidas repetidas, para identificá-los e agrupá-los para a formação de variedades mais produtivas e, simultaneamente, modelar possíveis oscilações da produção ao longo do tempo. A influência de ciclos produtivos foi considerada, tendo-se modelado as correlações existentes entre anos subsequentes, de forma eficiente, por meio da metodologia de análise de medidas repetidas, para avaliar produções dos materiais genéticos em cinco safras (1997 a 2001), o que possibilitou captar oscilações anuais na produção, ao longo do ciclo produtivo de *C. canephora*, relacionadas às maiores produções no terceiro e quarto anos. Além disso, a análise de grupo classificou os 50 materiais avaliados em apenas 10 grupos, de acordo com o padrão de oscilação da produção identificado pelo modelo de regressão adotado.

Termos para indexação: *Coffea canephora*, ciclo produtivo, melhoramento genético.

Repeated measure analysis in the clonal evaluation in 'Conilon' coffee

Abstract – The objective of this work was to evaluate the production of fifty clonal varieties coffee 'Conilon' (*Coffea canephora*), through the repeated measure methodology in order to identify and cluster them to obtain more productive genetic materials and to model, simultaneously, the correlation between subsequent years. The influence of productive cycles was efficiently regarded in the analysis by modeling the correlations between successive records of the genetic materials of five years (1997 to 2001), through this statistical methodology. The repeated measure analysis showed annual oscillations over the cycles of *C. canephora*, related to the largest productions in the third and fourth years. Moreover, the cluster analysis grouped the 50 clones into 10 groups, according to standard production oscillations identified through the adopted regression model.

Index terms: *Coffea canephora*, productive cycles, plant breeding.

Introdução

O agronegócio do café é uma das atividades mais importantes no aspecto socioeconômico no mundo. O café é cultivado em mais de 60 países, predominantemente nos menos desenvolvidos ou em desenvolvimento, e consumido prioritariamente por países ricos e desenvolvidos, que vêm cada vez mais exigindo produto de alta qualidade. Existem em torno de 80 espécies descritas no gênero *Coffea* (Mistro et al., 2004). Mas, da produção mundial de café, duas espécies, *Coffea canephora* e *C. arabica*, representam quase a totalidade comercializada (Ferreira et al., 2005).

Apesar de praticamente todo o café consumido no mundo ser originado de variedades de *C. arabica* e *C. canephora*, estas espécies são bem diferentes quanto ao sistema de reprodução, número de cromossomos, origens, ciclo, porte, tipos de grãos, exigências nutricionais, tolerância à seca e a pragas, doenças e nematóides, tipo e tamanho das cerejas e constituição bioquímica dos grãos, entre outras características. *C. canephora* tem despertado grande interesse, entre os melhoristas, por apresentar as características de alto potencial produtivo, grande rusticidade e maior quantidade de sólidos solúveis totais, em comparação a *C. arabica*. Essas características tornam a espécie muito desejada pelas indústrias e estimulam a adoção de

estratégias de melhoramento, com base na transferência de genes para outras espécies (Ferreira et al., 2005).

O sucesso do programa prático do melhoramento de espécies perenes depende, entre outros aspectos, do conhecimento do germoplasma disponível, da variação biológica entre espécies no gênero e entre populações dentro de espécies, e da variação entre indivíduos (Costa et al., 2005). Essas e outras informações, no contexto da genética quantitativa, possibilitam a determinação de metodologias de análises genético-biométricas adequadas.

Um importante aspecto biológico de espécies vegetais perenes é a presença de ciclos produtivos que tornam o melhoramento genético dessas espécies diferenciado de culturas anuais, por determinarem oscilações na produção ao longo do tempo. Assim, para que os resultados experimentais representem informações fidedignas da comparação entre diferentes materiais genéticos, a influência dos ciclos produtivos deve ser removida por meio de modelagem estatística. Uma forma eficiente de se realizar esta modelagem é por meio da análise de medidas repetidas (Littell et al., 1998), cuja utilização (no tempo), em cada indivíduo, é prática rotineira no melhoramento genético desses tipos de espécies (Resende et al., 2000).

A diferença entre a análise de medidas repetidas e a análise de parcela subdividida é que, na última, assume-se que as respostas de diferentes tempos, na mesma parcela, são igualmente correlacionadas. Na segunda, não é necessário assumir essa pressuposição, e é possível considerar, por exemplo, que as respostas de tempos mais próximos são, em geral, mais fortemente correlacionadas que as de tempos mais distantes (Littell et al., 1998).

Uma consequência imediata de se ignorarem diferentes correlações entre dados mensurados no tempo é que a significância aparente da diferença entre as médias dos tratamentos é grosseiramente exagerada, e a sensibilidade dos testes para interação é seriamente reduzida. De forma geral, a análise de medidas repetidas permite modelar o grau de homogeneidade das variâncias e covariâncias dos dados, nas diferentes épocas, por meio da utilização de matrizes específicas.

O objetivo deste trabalho foi analisar a produção de 50 clones de café 'Conilon', por meio da análise de medidas repetidas, para identificá-los e agrupá-los para a formação de variedades mais produtivas e, simultaneamente, modelar possíveis oscilações na produção ao longo do tempo.

Material e Métodos

Foram estudados 50 clones do Programa de Melhoramento Genético de Café 'Conilon', do Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural (Incaper), no Estado do Espírito Santo. O experimento foi instalado em março de 1993, na Fazenda Experimental de Sooretama, do Incaper, no Município de Sooretama, no Estado do Espírito Santo.

O local do experimento situa-se a 15°47'S, 43°18'W e à altitude de 40 m, em Latossolo Vermelho-Amarelo distrófico arenoso (80% de areia) de baixa fertilidade (Embrapa, 2006), com precipitação pluviométrica média anual de 1.200 mm, mal distribuída; a temperatura média anual é de 24°C, com umidade relativa média do ar de 80%; e topografia plana com vento sul predominante. O clima em janeiro, novembro e dezembro é úmido; em março, abril e outubro é parcialmente úmido; e em maio, junho, julho, agosto e setembro é seco.

O experimento foi implantado e conduzido sem irrigação, em delineamento experimental de blocos ao acaso, com 12 repetições. Cada parcela foi constituída por duas plantas, no espaçamento de 3 m entre linhas e 1,5 m entre plantas. O manejo, as adubações, os tratamentos culturais, as podas e as colheitas das plantas dos experimentos foram realizados conforme as recomendações técnicas da cultura, e as adubações foram feitas com base nos resultados das análises de solo do local. A produção foi avaliada em cinco safras, colhidas de 1997 a 2001.

Foi utilizado o seguinte modelo estatístico: $y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_k + \xi_{ijk}$, em que: y_{ijk} é a produção do clone i , no ano j , no bloco k ; μ é a média global; α_i é o efeito do clone i ; β_j é o efeito do ano j ; $(\alpha\beta)_{ij}$ é o efeito da interação entre o clone i e o ano j ; γ_k é o efeito do bloco k ; e ξ_{ijk} é o erro aleatório, $\xi_{ijk} \sim N(0, V)$, em que V é a matriz de covariância utilizada para modelar a dependência dos erros. Foi assumido que esta dependência dos erros ocorre entre a j -ésima e a m -ésima ($j \neq m$) avaliação anual de um mesmo clone, em um mesmo bloco e, assim, avaliaram-se as matrizes de covariância seguintes, correspondentes, respectivamente, aos componentes de variância, à simetria composta, à auto-regressiva de primeira ordem e à simetria composta heterogênea:

$$V = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$$

$$V = \begin{bmatrix} \sigma^2 + \sigma_1 & \sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma^2 + \sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma^2 + \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma_1 & \sigma^2 + \sigma_1 \end{bmatrix}$$

$$V = \begin{bmatrix} \sigma^2 & \rho\sigma^2 & \rho^2\sigma^2 & \rho^3\sigma^2 \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho^2\sigma^2 \\ \rho^2\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 \\ \rho^3\sigma^2 & \rho^2\sigma^2 & \rho\sigma^2 & \sigma^2 \end{bmatrix}$$

$$V = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho\sigma_1\sigma_3 & \rho\sigma_1\sigma_4 \\ \rho\sigma_2\sigma_1 & \sigma_2^2 & \rho\sigma_2\sigma_3 & \rho\sigma_2\sigma_4 \\ \rho\sigma_3\sigma_1 & \rho\sigma_3\sigma_2 & \sigma_3^2 & \rho\sigma_3\sigma_4 \\ \rho\sigma_4\sigma_1 & \rho\sigma_4\sigma_2 & \rho\sigma_4\sigma_3 & \sigma_4^2 \end{bmatrix}$$

Os termos σ^2 , σ e ρ correspondem, respectivamente, à variância, ao desvio-padrão e à correlação.

O modelo apresentado, considerando-se todas as matrizes de covariância utilizadas, foi ajustado aos dados pelo método da máxima verossimilhança, por meio do PROC MIXED (SAS Institute, 2002).

Para identificar qual foi a melhor matriz, utilizou-se o critério de informação de Akaike (AIC) (Akaike, 1974), conforme Littell et al. (1998), de forma que quanto menor o seu valor, melhor foi o ajuste do modelo em questão. Sua expressão é dada por: $AIC = -2\log \text{verossimilhança} + 2q$, em que: q é o número de parâmetros na matriz considerada.

Na análise da interação entre clones e tempo, ajustou-se um modelo polinomial cúbico, para descrever a produção em função do tempo, em cada clone. Embora, geralmente, a utilização deste grau polinomial seja criticado pela falta de interpretação biológica, neste trabalho ele é de grande utilidade, pois, no período considerado, observaram-se produções mínimas e máximas.

Fixou-se o efeito de clone para se aplicar a identidade de modelos, e agrupar curvas de produção estatisticamente iguais. Para isso, utilizou-se a análise de agrupamento, com o método centróide por meio do PROC CLUSTER

(SAS Institute, 2002), uma vez que o número de tratamentos era alto. Nesta análise, foram obtidos os valores dos desvios-padrão da raiz do quadrado médio (RMSSTD), em relação ao número de grupos (Khattree & Naik, 2000), que geraram o gráfico que disponibilizou a identificação do número ótimo de cluster em relação à máxima curvatura, observada geometricamente (Larson et al., 1998). Como complemento dessa análise, utilizou-se o PROC TREE (SAS Institute, 2002), para se visualizar o dendrograma e verificar quais clones pertenciam a cada grupo obtido pela discriminação estatística.

Ao se fixar o efeito de tempo, para comparar as produções médias dos clones dentro de cada ano, utilizou-se o teste de Scott-Knott, a 5% de probabilidade, que foi implementado no SAEG (SAEG, 2007) considerando o erro-padrão gerado pela análise de medidas repetidas efetuada no SAS.

Resultados e Discussão

Os valores do AIC obtidos: 9.975,6, 9.925,5, 9.922,0 e 9.221,5, respectivamente para as estruturas de matrizes componentes de variância, simetria composta, auto-regressiva de primeira ordem e simetria composta heterogênea, mostraram que esta última foi selecionada por ter apresentado melhor ajuste.

Uma vez selecionado o melhor modelo, procedeu-se ao desdobramento da interação tempo x clone. Pelo fator tempo, obteve-se maior variação de produções médias a partir do terceiro ano e, principalmente no quarto ano, quando a produção foi maior (Tabela 1).

Em todo o seu período de produção, a espécie apresenta ciclos de rendimento de quatro a sete anos, regulados por freqüentes podas de produção (Cilas et al., 2006). Neste trabalho, a primeira poda foi feita após a realização da quarta colheita (período recomendado para o Estado do Espírito Santo), com 12 mil ramos produtivos por hectare. Em razão de o cafeeiro 'Conilon' apresentar resposta muito favorável ao uso de tecnologias e efeito de bienalidade, que associado à poda e outros estresses pode levar a oscilações nas produções de ano para ano, justifica-se a queda de produção observada no quinto ano. Isto se dá principalmente em plantios não irrigados, como é o caso do trabalho em questão.

Ao se estudar o efeito de ano, estimou-se uma equação de regressão polinomial de terceiro grau, para cada um dos clones, no total de 50 equações. Portanto,

para se aplicar a análise de grupo e discriminar os modelos, é necessário que o número ótimo de grupos, ou seja, de equações estatisticamente diferentes, seja determinado.

A utilização de mais de dez grupos distintos não influencia de forma efetiva a diminuição da magnitude do RMSSTD (Figura 1 A). Esse valor ótimo é melhor observado na Figura 1 B, em que a maior distância

Tabela 1. Produção média de *Coffea canephora*, de 1997 a 2001, em quilogramas por parcela⁽¹⁾.

Clone	1º ano	2º ano	3º ano	4º ano	5º ano
1	1,3c	1,6b	4,8c	8,5e	2,5e
2	1,3c	0,5c	5,8b	10,6c	3,6d
3	2,2b	1,3b	7,0b	8,6e	5,4b
4	2,0b	0,6c	7,5a	5,6g	4,9c
5	1,3c	2,5b	4,9c	5,9g	3,9d
6	0,9c	1,1c	7,0b	9,5d	4,5c
7	1,0c	1,6b	6,6b	7,2f	2,8e
8	2,0b	1,3b	7,7a	10,6c	3,8d
9	1,0c	0,3c	2,8e	1,7j	3,1e
10	2,3b	0,6c	6,1b	10,0c	3,9d
11	1,1c	0,6c	4,1d	3,4i	4,1d
12	3,3a	1,8b	7,6a	10,4c	5,2b
13	1,9b	0,5c	5,3c	10,5c	4,7c
14	1,8b	0,3c	6,5b	7,0f	5,6b
15	2,5a	1,5b	5,4c	6,2g	3,9d
16	0,9c	0,2c	5,3c	5,0g	3,5d
17	1,7b	1,3b	6,6b	9,1d	2,8e
18	1,3c	0,4c	3,8d	5,3g	2,7e
19	3,3a	1,9b	6,7b	6,8f	3,9d
20	0,9c	1,4b	5,3c	4,9g	4,5c
21	1,4c	0,8c	4,6c	5,3g	4,6c
22	3,2a	1,7b	7,8a	11,9b	6,4a
23	1,1c	1,5b	5,0c	7,0f	6,5a
24	1,8b	1,0c	6,4b	7,6f	3,7d
25	1,1c	0,4c	2,5e	1,3j	4,5c
26	2,0b	0,5c	4,0d	6,0g	3,7d
27	1,3c	0,8c	5,7b	14,6a	3,0e
28	1,8b	1,7b	6,6b	10,0c	3,3d
29	0,5c	0,4c	2,8e	4,1h	3,3d
30	0,6c	0,7c	3,0e	5,9g	3,8d
31	3,5a	4,5a	8,2a	11,5b	4,1d
32	1,9b	1,7b	6,8b	9,2d	2,7e
33	1,4c	0,7c	5,9b	2,9i	2,9e
34	1,0c	0,7c	6,5b	8,5e	3,9d
35	0,5c	0,2c	3,7d	3,7i	2,5e
36	0,3c	0,1c	4,2c	4,5h	4,8c
37	2,7a	1,8b	8,4a	10,7c	2,6e
38	1,3c	0,5c	5,3c	8,1e	5,4b
39	0,4c	0,3c	4,7c	4,5h	2,6e
40	0,9c	0,8c	4,0d	6,2g	3,2e
41	1,3c	2,0b	6,2b	7,4f	4,2c
42	1,3c	0,4c	3,5d	5,2g	3,2d
43	0,5c	0,4c	4,5c	3,6i	2,9e
44	1,4c	0,6c	5,6b	5,9g	3,4d
45	1,6c	1,4b	6,4b	11,0c	5,2b
46	1,9b	0,4c	6,0b	6,4f	4,2c
47	2,6a	0,7c	4,6c	6,0g	3,5d
48	2,7a	2,4b	7,0b	10,6c	3,4d
49	2,7a	1,2b	6,5b	7,9e	5,2b
50	0,7c	0,8c	4,4c	5,6g	2,5e

⁽¹⁾Médias seguidas de letras iguais, na coluna, não diferem entre si pelo teste de Scott-Knott, a 5% de probabilidade.

corresponde à maior curvatura. O dendrograma que indica a formação de cada grupo está na Figura 2.

A classificação dos clones dentro de cada grupo foi a seguinte: 1) 40, 44, 42, 14, 46, 15, 18 e 19; 2) 16, 41, 30, 50, 4, 7 e 21; 3) 20, 33, 5, 43, 11 e 9; 4) 6, 34, 3, 38, 49, 26, 47, 1, 24 e 31; 5) 10, 22, 13, 2 e 37; 6) 29, 35, 23, 39 e 36; 7) 45, 48, 8 e 12; 8) 17, 28 e 32; 9) 25; 10) 27.

A representação da curva média da produção de cada grupo está na Figura 3. No quarto ano, que corresponde biologicamente ao ano de maior produção, a variação das produções entre os grupos distintos foi mais acentuada que em outros pontos da curva, o que conduz aos mesmos resultados obtidos pelos testes de médias da Tabela 1. Esses valores podem ter ocorrido em razão da grande variabilidade genética entre os clones e das respostas diferenciadas dos materiais genéticos ao efeito temporal ano e aos vários estresses, sobretudo o hídrico, uma vez que o trabalho foi implantado em local arenoso, com vento intenso, seco e com temperatura elevada, sem irrigação e com má distribuição de chuva. Resultados semelhantes na variação de produção e presença de grande variabilidade genética, em clones de *C. canephora*, foram observados nos trabalhos de Bragança et al. (2001) e (Ferrão, 2004).

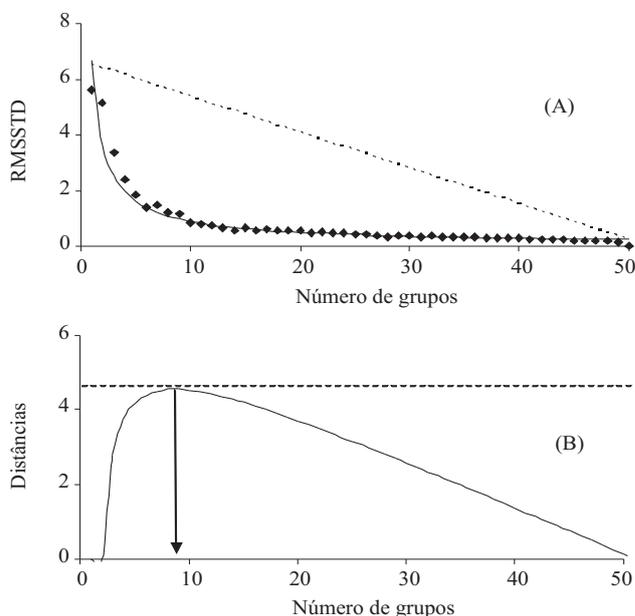


Figura 1. Gráfico dos desvios-padrão da raiz do quadrado médio (RMSSTD) em relação ao número de grupos (A); método da máxima curvatura geométrica, para identificar o número de grupos (B).

Os trabalhos de divergência genética com estudos de agrupamentos são de primordial importância no planejamento de programas direcionados à obtenção de

híbridos heteróticos, na formação de populações-base para futuros programas de melhoramento e, também, na seleção de materiais genéticos para a formação de



Figura 2. Dendrograma indicador dos grupos formados, em relação às equações polinomiais estimadas.

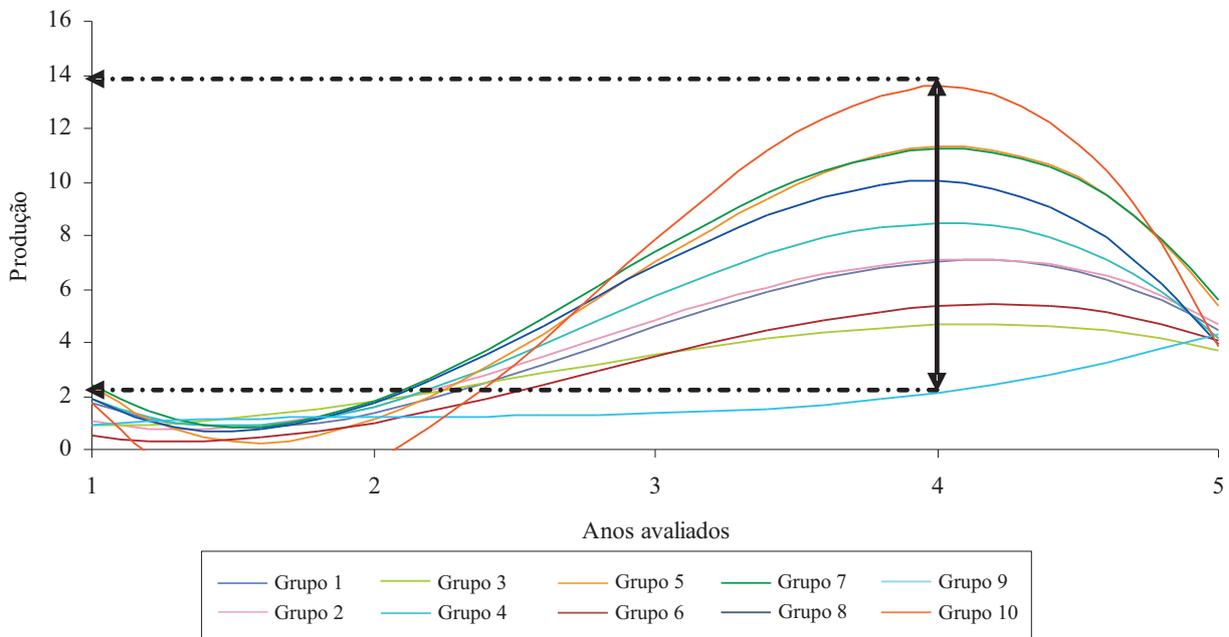


Figura 3. Curvas de produção média (quilograma por parcela) dos grupos de clones de *Coffea canephora* avaliados de 1997 a 2001.

variedades sintéticas. Assim, na seleção de progenitores de café 'Conilon', devem-se levar em consideração a associação entre estudos de divergência genética e a produtividade por ciclo.

Conclusões

1. A análise de medidas repetidas possibilita captar oscilações anuais na produção, ao longo do ciclo produtivo do *Coffea canephora*.

2. A análise de grupos classifica os clones em grupos menores, de acordo com o padrão de produção identificado pela análise de regressão adotada.

Referências

- AKAIKE, H. A new look at the statistical model identification. **IEEE transactions on Automatic Control**, v.19, p.716-723, 1974.
- BRAGANÇA, S.M.; CARVALHO, C.H.S. de; FONSECA, A.F.A. da; FERRÃO, R.G. Variedades clonais de café 'Conilon' para o Estado do Espírito Santo. **Pesquisa Agropecuária Brasileira**, v.36, p.765-770, 2001.
- CILAS, C.; BAR-HEN, A.; MONTAGNON, C.; GODIN, C. Definition of architectural ideotypes for good yield capacity in *Coffea canephora*. **Annals of Botany**, v.97, p.405-411, 2006.
- COSTA, R.B. da; RESENDE, M.D.V. de; CONTINI, A.Z.; REGO, F.L.H.; ROA, R.A.R.; MARTINS, W.J. Avaliação genética de indivíduos de erva-mate (*Ilex paraguariensis* St. Hil.) na região de Caarapó, MS, pelo procedimento REML/BLUP. **Ciência Florestal**, v.15, p.371-376, 2005.
- EMBRAPA. **Sistema brasileiro de classificação de solos**. 2.ed. Rio de Janeiro: Embrapa Solos, 2006. 306p.
- FERRÃO, R.G. **Biometria aplicada ao melhoramento genético do café 'Conilon'**. 2004. 256p. Tese (Doutorado) - Universidade Federal de Viçosa, Viçosa.
- FERREIRA, A.; CECON, P.R.; CRUZ, C.D.; FERRÃO, R.G.; SILVA, M.F. da; FONSECA, A.F.A. da; FERRÃO, M.A.G. Seleção simultânea de *Coffea canephora* por meio da combinação de análise de fatores e índices de seleção. **Pesquisa Agropecuária Brasileira**, v.40, p.1189-1195, 2005.
- KHATTREE, R.; NAIK, D. **Multivariate data reduction and discrimination with SAS software**. Cary: SAS Institute, 2000. 574p.
- LARSON, R.; HOSTETLER, R.; EDWARDS, B. **Cálculo com aplicações**. 5.ed. Rio de Janeiro: LTC, 1998.
- LITTELL, R.C.; HENRY, P.R.; AMMERMAN, C.B. Statistical analysis of repeated measures using SAS procedures. **Journal of Animal Science**, v.76, p.1216-1231, 1998.
- MISTRO, J.C.; FAZUOLI, L.C.; GONÇALVES, P. de S.; GUERREIRO FILHO, O. Estimates of genetic parameters and expected genetic gains with selection in robust coffee. **Crop Breeding and Applied Biotechnology**, v.4, p.86-91, 2004.
- RESENDE, M.D.V. de; SIMEÃO, R.M.; FERNANDES, J.S.C. BLUP individual com medidas repetidas aplicado ao melhoramento de espécies perenes. **Ciência e Agrotecnologia**, v.24, p.986-997, 2000.
- SAEG. **SAEG: sistema para análises estatísticas**, versão 9.1. Viçosa: UFV, 2007.
- SAS INSTITUTE. **SAS user's guide: statistics**, version 9.1. Cary: SAS Institute, 2002.

Recebido em 14 de dezembro de 2007 e aprovado em 6 de agosto de 2008